```
   1 TATCCGCACCTCCGCTCCCACCCGGCGCCTCGGCGCGCCCGCCCTCCGATGCGCTCAGCG
   1 ───────────────────────▶                              M  R  S  A
          F-1006

  61 GCCGCAGCTCCTCGGAGTCCCGCGGTGGCCACCGAGTCTCGCCGCTTCGCCGCAGCCAGG
   5  A  A  A  P  R  S  P  A  V  A  T  E  S  R  R  F  A  A  A  R

 121 TGGCCCGGGTGGCGCTCGCTCCAGCGGCCGGCGCGGCGGAGCGGGCGGGGCGGCGGTGGC
  25  W  P  G  W  R  S  L  Q  R  P  A  R  R  S  G  R  G  G  G  G

 181 GCGGCCCCGGGACCGTATCCCTCCGCCGCCCCTCCCCCGCCCGGCCCCGGCCCCCCTCCC
  45  A  A  P  G  P  Y  P  S  A  A  P  P  P  P  G  P  G  P  P  P

 241 TCCCGGCAGAGCTCGCCTCCCTCCGCCTCAGACTGTTTTGGTAGCAACGGCAACGGCGGC
  65  S  R  Q  S  S  P  P  S  A  S  D  C  F  G  S  N  G  N  G  G

 301 GGCGCGTTTCGGCCCGGCTCCCGGCGGCTCCTTGGTCTCGGCGGGCCTCCCCGCCCCTTC
                        ──────────R-1002
  85  G  A  F  R  P  G  S  R  R  L  L  G  L  G  G  P  P  R  P  F

 361 GTCGTCGTCCTTCTCCCCCTCGCCAGCCCGGGCGCCCCTCCGGCCGCGCCAACCCGCGCC
 105  V  V  V  L  L  P  L  A  S  P  G  A  P  P  A  A  P  T  R  A  .

 421 TCCCCGCTCGGCGCCCGTGCGTCCCCGCCGCGTTCCGGCGTCTCCTTGGCGCGCCCGGCT
 125  S  P  L  G  A  R  A  P  P  P  R  S  G  V  S  L  A  R  P  A

 481 CCCGGCTGTCCCCGCCCGGCGTGCGAGCCGGTGTATGGGCCCCTCACCATGTCGCTGAAG
                                        ──────F-1─────
 145  P  G  C  P  R  P  A  C  E  P  V  Y  G  P  L  T  M  S  L  K

 541 CCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAG
 165  P  Q  Q  Q  Q  Q  Q  Q  Q  Q  Q  Q  Q  Q  Q  Q  Q  Q  Q  Q

 601 CAGCAGCAGCAGCCGCCGCCCGCGGCTGCCAATGTCCGCAAGCCCGGCGGCAGCGGCCTT
                                    ─────────R-2──────◀
 185  Q  Q  Q  Q  P  P  P  A  A  A  N  V  R  K  P  G  G  S  G  L

 661 CTAGCGTCGCCCGCCGCCGCGCCTTCGCCGTCCTCGTCCTCGGTCTCCTCGTCCTCGGCC
     ──────────R-1
 205  L  A  S  P  A  A  A  P  S  P  S  S  S  S  V  S  S  S  S  A

 721 ACGGCTCCCTCCTCGGTGGTCGCGGCGACCTCCGGCGGCGGGAGGCCCGGCCTGGGCAGA
 225  T  A  P  S  S  V  V  A  A  T  S  G  G  G  R  P  G  L  G  R

 781 GGTCGAAACAGTAACAAAGGACTGCCTCAGTCTACGATTTCTTTTGATGGAATCTATGCA
 245  G  R  N  S  N  K  G  L  P  Q  S  T  I  S  F  D  G  I  Y  A

 841 AATATGAGGATGGTTCATATACTTACATCAGTTGTTGGCTCCAAATGTGAAGTACAAGTG
 265  N  M  R  M  V  H  I  L  T  S  V  V  G  S  K  C  E  V  Q  V

 901 AAAAATGGAGGTATATATGAAGGAGTTTTTAAAACTTACAGTCCGAAGTGTGATTTGGTA
 285  K  N  G  G  I  Y  E  G  V  F  K  T  Y  S  P  K  C  D  L  V

 961 CTTGATGCCGCACATGAGAAAAGTACAGAATCCAGTTCGGGGCCGAAACGTGAAGAAATA
 305  L  D  A  A  H  E  K  S  T  E  S  S  S  G  P  K  R  E  E  I

1021 ATGGAGAGTATTTTTGTTCAAATGTTCAGACTTTGTTGTGGTACAGTTTAAAGATATGGAC
 325  M  E  S  I  L  F  K  C  S  D  F  V  V  V  Q  F  K  D  M  D

1081 TCCAGTTATGCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGCTAAAGTGAATGGC
 345  S  S  Y  A  K  R  D  A  F  T  D  S  A  I  S  A  K  V  N  G

1141 GAACACAAAGAGAAGGACCTGGAGCCCTGGGATGCAGGTGAACTCACAGCCAATGAGGAA
 365  E  H  K  E  K  D  L  E  P  W  D  A  G  E  L  T  A  N  E  E

1201 CTTGAGGCTTTGGAAAATGACGTATCTAATGGATGGGATCCCAATGATATGTTTCGATAT
 385  L  E  A  L  E  N  D  V  S  N  G  W  D  P  N  D  M  F  R  Y

1261 AATGAAGAAAATTATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTG
 405  N  E  E  N  Y  G  V  V  S  T  Y  D  S  S  L  S  S  Y  T  V
```

Fig. 1

```
1321 CCCTTAGAAAGAGATAACTCAGAAGAATTTTTAAAACGGGAAGCAAGGGCAAACCAGTTA
 425  P   L   E   R   D   N   S   E   E   F   L   K   R   E   A   R   A   N   Q   L

1381 GCAGAAGAAATTGAGTCAAGTGCCCAGTACAAAGCTCGAGTGGCCCTGGAAAACGATGAT
 445  A   E   E   I   E   S   S   A   Q   Y   K   A   R   V   A   L   E   N   D   D

1441 AGGAGTGAGGAAGAAAAATACACAGCAGTTCAGAGAAATTCCAGTGAACGTGAGGGGCAC
 465  R   S   E   E   E   K   Y   T   A   V   Q   R   N   S   S   E   R   E   G   H

1501 AGCATAAACACTAGGGAAAATAAATATATTCCTCCTGGACAAAGAAATAGAGAAGTCATA
 485  S   I   N   T   R   E   N   K   Y   I   P   P   G   Q   R   N   R   E   V   I

1561 TCCTGGGGAAGTGGGAGACAGAATTCACCGCGTATGGGCCAGCCTGGATCGGGCTCCATG
 505  S   W   G   S   G   R   Q   N   S   P   R   M   G   Q   P   G   S   G   S   M

1621 CCATCAAGATCCACTTCTCACACTTCAGATTTCAACCCGAATTCTGGTTCAGACCAAAGA
 525  P   S   R   S   T   S   H   T   S   D   F   N   P   N   S   G   S   D   Q   R

1681 GTAGTTAATGGAGGTGTTCCCTGGCCATCGCCTTGCCCATCTCCTTCCTCTCGCCCACCT
 545  V   V   N   G   G   V   P   W   P   S   P   C   P   S   P   S   S   R   P   P

1741 TCTCGCTACCAGTCAGGTCCCAACTCTCTTCCACCTCGGGCAGCCACCCCTACACGGCCG
 565  S   R   Y   Q   S   G   P   N   S   L   P   P   R   A   A   T   P   T   R   P

1801 CCCTCCAGGCCCCCCTCGCGGCCATCCAGACCCCCGTCTCACCCCTCTGCTCATGGTTCT
 585  P   S   R   P   P   S   R   P   S   R   P   P   S   H   P   S   A   H   G   S

1861 CCAGCTCCTGTCTCTACTATGCCTAAACGCATGTCTTCAGAAGGGCCTCCAAGGATGTCC
 605  P   A   P   V   S   T   M   P   K   R   M   S   S   E   G   P   P   R   M   S

1921 CCAAAGGCCCAGCGACATCCTCGAAATCACAGAGTTTCTGCTGGGAGGGGTTCCATATCC
 625  P   K   A   Q   R   H   P   R   N   H   R   V   S   A   G   R   G   S   I   S

1981 AGTGGCCTAGAATTTGTATCCCACAACCCACCCAGTGAAGCAGCTACTCCTCCAGTAGCA
 645  S   G   L   E   F   V   S   H   N   P   P   S   E   A   A   T   P   P   V   A

2041 AGGACCAGTCCCTCGGGGGGAACGTGGTCATCAGTGGTCAGTGGGGTTCCAAGATTATCC
 665  R   T   S   P   S   G   G   T   W   S   S   V   V   S   G   V   P   R   L   S

2101 CCTAAAACTCATAGACCCAGGTCTCCCAGACAGAACAGTATTGGAAATACCCCCAGTGGG
 685  P   K   T   H   R   P   R   S   P   R   Q   N   S   I   G   N   T   P   S   G

2161 CCAGTTCTTGCTTCTCCCCAAGCTGGTATTATTCCAACTGAAGCTGTTGCCATGCCTATT
 705  P   V   L   A   S   P   Q   A   G   I   I   P   T   E   A   V   A   M   P   I

2221 CCAGCTGCATCTCCTACGCCTGCTAGTCCTGCATCGAACAGAGCTGTTACCCCTTCTAGT
 725  P   A   A   S   P   T   P   A   S   P   A   S   N   R   A   V   T   P   S   S

2281 GAGGCTAAAGATTCCAGGCTTCAAGATCAGAGGCAGAACTCTCCTGCAGGGAATAAAGAA
 745  E   A   K   D   S   R   L   Q   D   Q   R   Q   N   S   P   A   G   N   K   E

2341 AATATTAAACCCAATGAAACATCACCTAGCTTCTCAAAAGCTGAAAACAAAGGTATATCA
 765  N   I   K   P   N   E   T   S   P   S   F   S   K   A   E   N   K   G   I   S

2401 CCAGTTGTTTCTGAACATAGAAAACAGATTGATGATTTAAAGAAATTTAAGAATGATTTT
 785  P   V   V   S   E   H   R   K   Q   I   D   D   L   K   K   F   K   N   D   F

2461 AGGTTACAGCCAAGTTCTACTTCTGAATCTATGGATCAACTACTAAACAAAAATAGAGAG
 805  R   L   Q   P   S   S   T   S   E   S   M   D   Q   L   L   N   K   N   R   E

2521 GGAGAAAAATCAAGAGATTTGATCAAAGACAAAATTGAACCAAGTGCTAAGGATTCTTTC
 825  G   E   K   S   R   D   L   I   K   D   K   I   E   P   S   A   K   D   S   F

2581 ATTGAAAATAGCAGCAGCAACTGTACCAGTGGCAGCAGCAAGCCGAATAGCCCCAGCATT
 845  I   E   N   S   S   S   N   C   T   S   G   S   S   K   P   N   S   P   S   I
```

Fig. 2

2641 TCCCCTTCAATACTTAGTAACACGGAGCACAAGAGGGGACCTGAGGTCACTTCCCAAGGG
865 S   P   S   I   L   S   N   T   E   H   K   R   G   P   E   V   T   S   Q   G

2701 GTTCAGACTTCCAGCCCAGCATGTAAACAAGAGAAAGACGATAAGGAAGAGAAGAAAGAC
885 V   Q   T   S   S   P   A   C   K   Q   E   K   D   D   K   E   E   K   K   D

2761 GCAGCTGAGCAAGTTAGGAAATCAACATTGAATCCCAATGCAAAGGAGTTCAACCCACGT
905 A   A   E   Q   V   R   K   S   T   L   N   P   N   A   K   E   F   N   P   R

2821 TCCTTCTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACCTCGGCCTCAAGCACAACCT
F-13
925 S   F   S   Q   P   K   P   S   T   T   P   T   S   P   R   P   Q   A   Q   P

2881 AGCCCATCTATGGTGGGTCATCAACAGCCAACTCCAGTTTATACTCAGCCTGTTTGTTTT
945 S   P   S   M   V   G   H   Q   Q   P   T   P   V   Y   T   Q   P   V   C   F

2941 GCACCAAATATGATGTATCCAGTCCCAGTGAGCCCAGGCGTGCAACCTTTATACCCAATA
965 A   P   N   M   M   Y   P   V   P   V   S   P   G   V   Q   P   L   Y   P   I

3001 CCTATGACGCCCATGCCAGTGAATCAAGCCAAGACATATAGAGCAGTACCAAATATGCCC
985 P   M   T   P   M   P   V   N   Q   A   K   T   Y   R   A   V   P   N   M   P

3061 CAACAGCGGCAAGACCAGCATCATCAGAGTGCCATGATGCACCCAGCGTCAGCAGCGGGC
1005 Q   Q   R   Q   D   Q   H   H   Q   S   A   M   M   H   P   A   S   A   A   G

3121 CCACCGATTGCAGCCACCCCACCAGCTTACTCCACGCAATATGTTGCCTACAGTCCTCAG
1025 P   P   I   A   A   T   P   P   A   Y   S   T   Q   Y   V   A   Y   S   P   Q

3181 CAGTTCCCAAATCAGCCCCTTGTTCAGCATGTGCCACATTATCAGTCTCAGCATCCTCAT
1045 Q   F   P   N   Q   P   L   V   Q   H   V   P   H   Y   Q   S   Q   H   P   H

3241 GTCTATAGTCCTGTAATACAGGGTAATGCTAGAATGATGGCACCACCAACACACGCCCAG
1065 V   Y   S   P   V   I   Q   G   N   A   R   M   M   A   P   P   T   H   A   Q

3301 CCTGGTTTAGTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGACGCATGCGATG
1085 P   G   L   V   S   S   S   A   T   Q   Y   G   A   H   E   Q   T   H   A   M

3361 TATGCATGTCCCAAATTACCATACAACAAGGAGACAAGCCCTTCTTTCTACTTTGCCATT
1105 Y   A   C   P   K   L   P   Y   N   K   E   T   S   P   S   F   Y   F   A   I

3421 TCCACGGGCTCCCCTTGCTCAGCAGTATGCGCACCCTAACGCTACCCTGCACCCACATACT
1125 S   T   G   S   L   A   Q   Q   Y   A   H   P   N   A   T   L   H   P   H   T

3481 CCACACCCTCAGCCTTCAGCTACCCCCACTGGACAGCAGCAAAGCCAACATGGTGGAAGT
1145 P   H   P   Q   P   S   A   T   P   T   G   Q   Q   Q   S   Q   H   G   G   S

3541 CATCCTGCACCCAGTCCTGTTCAGCACCATCAGCACCAGGCCGCCCAGGCTCTCCATCTG
1165 H   P   A   P   S   P   V   Q   H   H   Q   H   Q   A   A   Q   A   L   H   L

3601 GCCAGTCCACAGCAGCAGTCAGCCATTTACCACGCGGGGCTTGCGCCAACTCCACCCTCC
1185 A   S   P   Q   Q   Q   S   A   I   Y   H   A   G   L   A   P   T   P   P   S

3661 ATGACACCTGCCTCCAACACGCAGTCGCCACAGAATAGTTTCCCAGCAGCACAACAGACT
1205 M   T   P   A   S   N   T   Q   S   P   Q   N   S   F   P   A   A   Q   Q   T

3721 GTCTTTACGATCCATCCTTCTCACGTTCAGCCGGCGTATACCAACCCACCCCACATGGCC
1225 V   F   T   I   H   P   S   H   V   Q   P   A   Y   T   N   P   P   H   M   A

3781 CACGTACCTCAGGCTCATGTACAGTCAGGAATGGTTCCTTCTCATCCAACTGCCCATGCG
1245 H   V   P   Q   A   H   V   Q   S   G   M   V   P   S   H   P   T   A   H   A

3841 CCAATGATGCTAATGACGACACAGCCACCCGGCGGTCCCCAGGCCGCCCTCGCTCAAAGT
1265 P   M   M   L   M   T   T   Q   P   P   G   G   P   Q   A   A   L   A   Q   S

3901 GCACTACAGCCCATTCCAGTCTCGACAACAGCGCATTTCCCCTATATGACGCACCCTTCA
1285 A   L   Q   P   I   P   V   S   T   T   A   H   F   P   Y   M   T   H   P   S

3961 GTACAAGCCCACCACCAACAGCAGTTGTAAGGCTGCCCTGGAGGAACCGAAAGGCCAAAT
1305 V   Q   A   H   H   Q   Q   Q   L   *

Fig. 3

```
4021  TCCCTCCTCCCTTCTACTGCTTCTACCAACTGGAAGCACAGAAAACTAGAATTTCATTTA
4081  TTTTGTTTTTAAAATATATATGTTGATTTCTTGTAACATCCAATAGGAATGCTAACAGTT
4141  CACTTGCAGTGGAAGATACTTGGACCGAGTAGAGGCATTTAGGAACTTGGGGGCTATTCC
4201  ATAATTCCATATGCTGTTTCAGAGTCCCGCAGGTACCCCAGCTCTGCTTGCCGAAACTGG
4261  AAGTTATTTATTTTTTAATAACCCTTGAAAGTCATGAACACATCAGCTAGCAAAAGAAGT
4321  AACAAGAGTGATTCTTGCTGCTATTACTGCT(A)$_n$
```
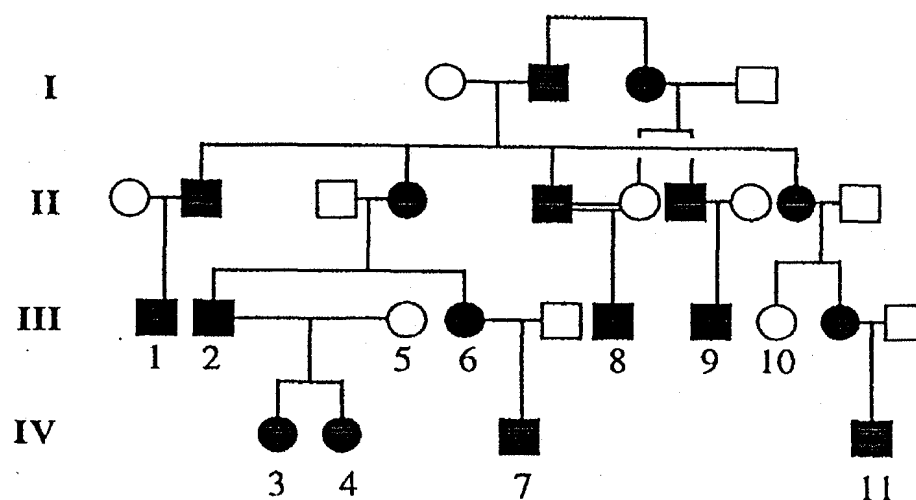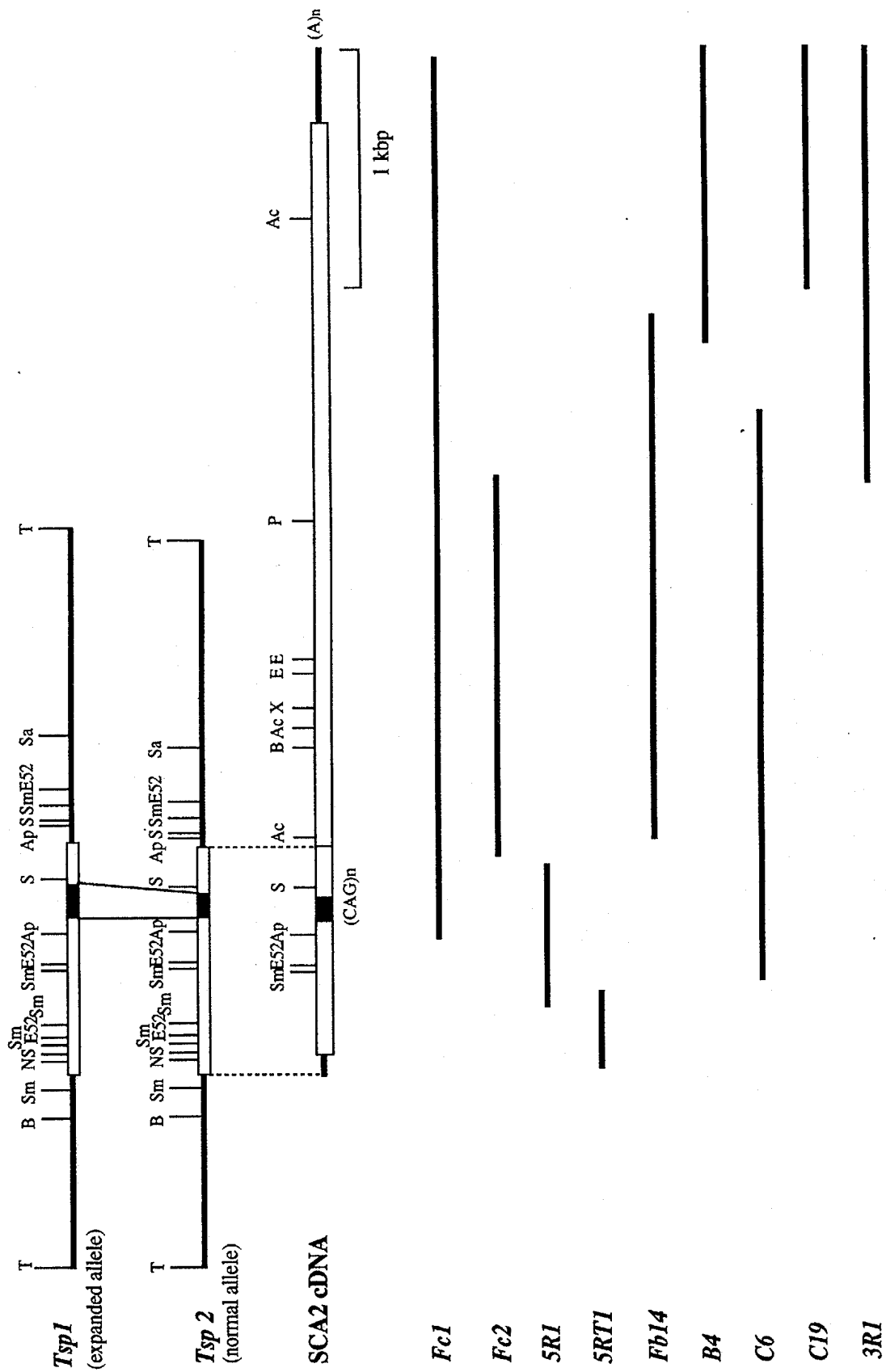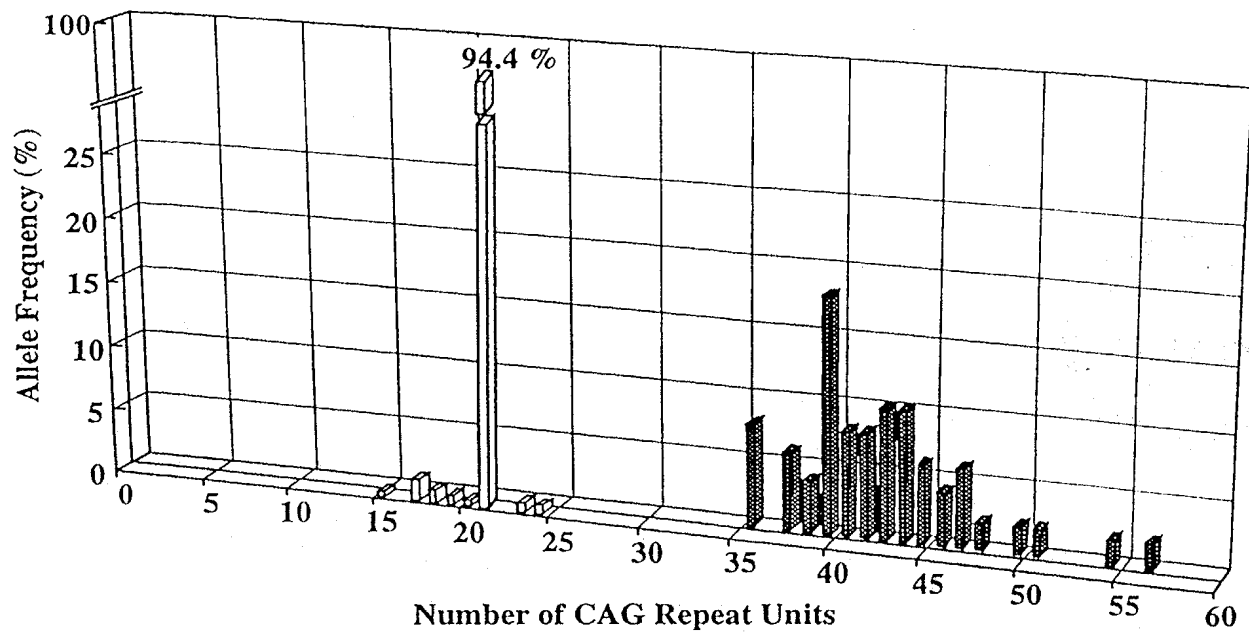
Fig. 4

Fig. 5

Fig. 6

09/101132

Fig. 7